DOI: 10.1002/cbic.200800314

The Entropy Balance of Nostocyclopeptide Macrocyclization Analysed by NMR Spectroscopy

Sebastian Enck,[a] Florian Kopp,[a, b] Mohamed A. Marahiel,*[a] and Armin Geyer*[a]

Nature has developed complex enzymatic machineries that promote the macrocyclization of linear precursors and provide a multitude of macrocyclic natural polyketides, peptides, and depsipeptides.[1,2] Macrocyclization of a ligand generally scales its receptor affinity and selectivity.[3–5] The large number of biologically active macrocyclic agents found in nature has prompted chemists to develop synthetic macrocyclization strategies for a large number of valuable protein ligands.[6,7] Synthetic cyclic peptides like the anti-cancer agents octreotide[8] and cilengitide[9] are success stories of drug design, and their potencies are ascribed to their bioactive conformations.[10] The modular composition of peptides facilitates the systematic variation of side chains, chirality and backbone chain length, and expedites the analysis of how these parameters affect biological activity. As determinants of the three-dimensional structure, these factors should also influence the inclination of a linear precursor to form a ring-closed structure. However, the thermodynamics of the ring-chain equilibrium of complex biomolecules has so far resisted quantitative experimental analysis, and therefore medicinal chemistry relies on computer modeling.[11] Through the establishment of the temperature dependent ring-chain equilibria of imino peptides, we were able to experimentally characterise the macrocyclization of a biomolecule and quantify the involved reaction entropy. Although obtained for a single ring size, these results are instructive for macrocyclization in general.

Biosynthetic macrocyclizations, like cyclopeptide formation, are generally irreversible in aqueous solution—a feature which adds to the stability of the macrocycle but prevents formation of the ring-chain equilibrium necessary for thermodynamic analysis (Figure 1 A). In this context, the nostocyclopeptides A1 (ncpA1) and A2 (ncpA2) turned out to be suitable candidates for the investigation of the macrocyclization process (Figure 1 C). These cyclic heptapeptides stand out by a hitherto unique backbone imino linkage, which is formed between a C-terminal aldehyde hydrate and an N-terminal amine (Figure 1 B).[12–14] The reversibility of imine formation in aqueous solution[15,16] should, under appropriate conditions, allow for the thermodynamic analysis of the equilibria between the linear and the cyclic peptides.

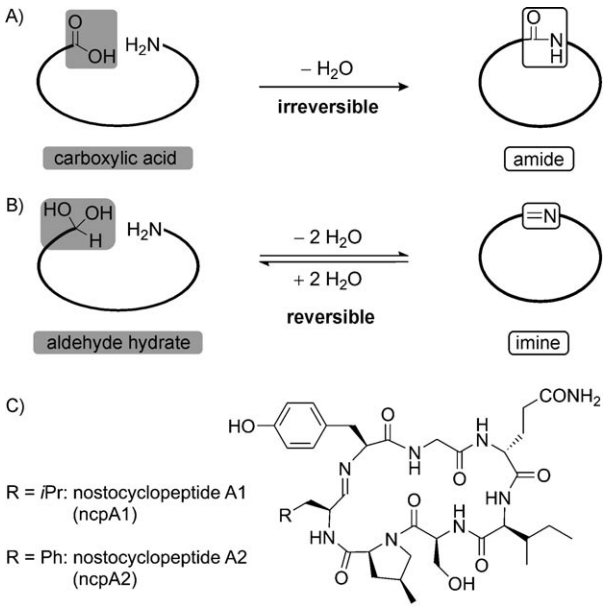

Figure 1. A) Cyclopeptides are ring-closed by condensation of a C-terminal carboxylic acid (or an activated analog) and an N-terminal amine, yielding a macrocyclic lactam. B) On the other hand, imines rarely occur as structural elements because of the reversibility of their formation. Compared to lactam formation, imine ring closure is entropically more favoured, as two water molecules instead of one are released. C) The two nostocyclopeptides, isolated from the terrestrial cyanobacterium *Nostoc* sp. ATCC 53789, represent the only known cyclic peptides which contain an imino linkage. Both peptides differ in the C-terminal amino acid, which is Leu in the case of nostocyclopeptide A1 (ncpA1), and Phe in the case of nostocyclopeptide A2 (ncpA2). Both also contain a D-configurated Gln and the nonproteinogenic amino acid (2*S*,4*S*)-methylproline.

Nostocyclopeptides have all the prerequisites necessary for the quantitative examination of a ring-chain equilibrium, including great selectivity for head-to-tail intramolecular cyclization and reluctance to form dimers or other linear or cyclic oligomerization byproducts. This is remarkable, as intermolecular oligomerization always competes with intramolecular ring closure[17,18] and normally occurs for macrocyclic imines, like the numerous examples studied in the field of supramolecular chemistry.[19] In contrast, the nostocyclopeptides, though highly decorated with stereocenters and functional groups, exhibit simple ring-chain equilibria that are concentration independent in the range studied. A second peculiarity of the nostocyclopeptide ring-chain tautomerism is the stereochemical self-cleaning that accompanies cyclic imine formation. The linear aldehyde hydrates appear as epimeric mixtures of (*R*)- and (*S*)-Leu in the case of ncpA1, and of (*R*)-Phe and (*S*)-Phe in the case of ncpA2. Yet, the cyclization process always results in stereochemically pure cyclopeptides.

[a] *S. Enck, Dr. F. Kopp, Prof. Dr. M. A. Marahiel, Prof. Dr. A. Geyer*
Fachbereich Chemie, Philipps-Universität-Marburg
Hans-Meerwein-Strasse, 35032 Marburg (Germany)
Fax: (+49) 6421-2822021
E-mail: geyer@staff.uni-marburg.de

[b] *Dr. F. Kopp*
Present address: The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

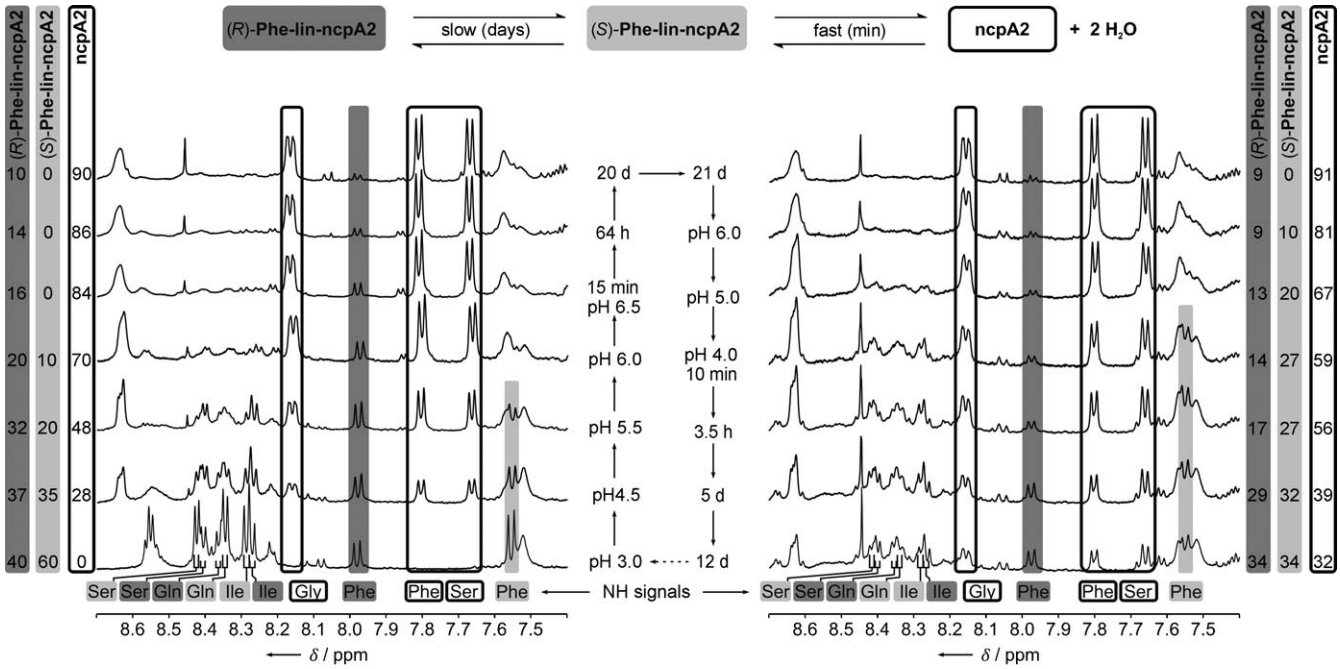

Figure 2. ^{1}H NMR spectra in the range of the amide protons were recorded from the ncpA2 sample at different pH values and timepoints using the Water suppression by Gradient-Tailored Excitation (Watergate) method. (Spectra were recorded at 300 K in a 5:1 solution of H_2O/D_2O with a KH_2PO_4/H_3PO_4 buffer solution). The pH value was increased by adding solid Na_2CO_3 and subsequently decreased by adding 0.01 M aqueous H_3PO_4. Signals resulting from the linear peptide aldehyde hydrate with (*R*)-Phe ((*R*)-Phe-lin-ncpA2) are highlighted in dark grey, while signals resulting from the epimeric (*S*)-Phe-lin-ncpA2 are highlighted in light grey. The cyclopeptide signals (ncpA2, with (*S*)-Phe) are framed. The change of relative signal intensities (listed as percentages) shows the existence of two equilibria, which are reached on minute (cyclization) and day timescales (epimerization). The self-purifying behaviour is visible from the fast decrease in the amount of *R* epimer, as the cyclization of the *S* epimer occurs much faster than the epimerization of *R* to *S* epimer. An alternative cyclization route (formation of *R* macrocycle from linear *R* epimer and subsequent epimerization to *S* macrocycle) can be excluded, as no further imine is visible in the NMR spectra. At pH 6.5 the sample was allowed to equilibrate over several days. The sample slowly underwent epimerization of the *R* to *S* epimer, which finally yielded more macrocycle. By subsequent acidification, the reverse process was monitored. The ring-opening first yields more linear *S* epimer, while the slower epimerization process further proceeds over several days after pH change.

^{1}H NMR spectroscopy can be used to quantitatively analyse the ratios of all contributing ring-chain isomers[20] and therefore allows the direct monitoring of the nostocyclopeptide equilibria, which can be controlled by changing pH and temperature. The linear aldehyde hydrates that dominate at pH 3 cyclize at neutral pH within a few minutes without the need of an activating reagent.[14] In an intermediate pH range, both species coexist, and the ratios of linear and cyclic peptides can be determined by comparing the corresponding NMR signal integrals. The ^{1}H NMR spectra depicted in Figure 2 illustrate the pH dependence of ncpA2 cyclization and epimerization.

As the remarkable cyclization behaviour of both nostocyclopeptides was expected to correlate with their amino acid composition, we elucidated their relative orientations and local mobilities by NMR spectroscopy and determined the solution conformations of ncpA2 (Figure 3A). As the naturally occurring (2*S*,4*S*)-methylproline had been replaced by proline for synthetic reasons, we also used proline in the structural calculations. A comparison of the geometries of linear and cyclic species revealed that an effective cyclization process relies on an appropriate combination of preorganized[17] backbone areas (characterised as non-averaged *J* couplings and NOEs) and predisposed[17] segments (visible as mainly averaged NMR parameters), as illustrated in Figure 3B. In the linear peptide, the Gly acts like a hinge and allows the N-terminal Tyr to snap into the emerging macrocycle. The preorientation of the linear precursors also manifests itself in the preformed β-turn as well as in the Pro amides, which exist to an unusually large extent (ca. 90%) in the *trans* conformation. The depicted backbone geometries also apply for the respective species of ncpA1, as no noticeable differences in the NMR parameters were obtained. However, the differences between ncpA1 and ncpA2 with respect to the C-terminal segments turned out to be far from negligible. Both peptides differ only in the C-terminal amino acid, which is an aliphatic Leu in ncpA1 and an aromatic Phe in ncpA2.

This subtle variation, however, has a considerable effect on the side chain orientations and mobilities as visible from *J* coupling[21] and NOEs. Figure 4A shows the different hydrophobic side chain interactions in the terminal segments according to the side chain variation before and after cyclization. Looking at the differences between the ncpA1 and ncpA2 species, the question arises: how much influence do noncovalent interactions have on the change in entropy that occurs during macrocyclization?

As seen in Figure 4A, NMR parameters give a detailed picture of the local relative orientations and dynamics of side chains. Yet, this data alone cannot provide the desired thermodynamic data of macrocyclization, which affects the global mobility of the peptide by linking both backbone ends; it is the

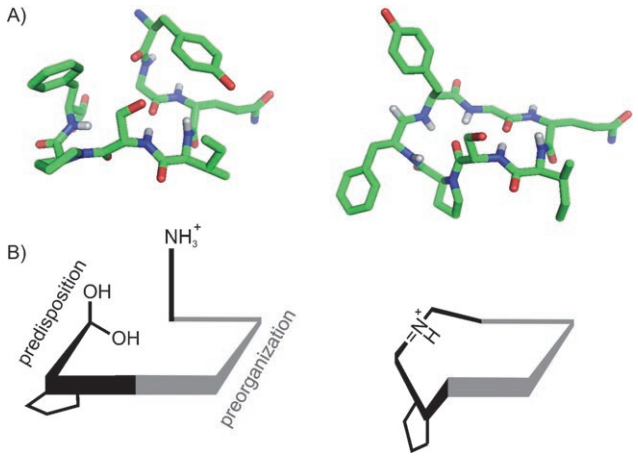

Figure 3. A) Average structures of Pro-γ-desmethyl-ncpA2 before (left) and after (right) ring closure, calculated from NMR parameters. Carbons are shown in green, oxygens in red, and nitrogens in blue. For clarity, only the backbone protons are depicted (in white). B) Schematic illustration of the peptide backbone, Pro residue and reacting functional groups of linear and cyclic ncpA2. The changes in backbone geometry upon cyclization give evidence of two distinguishable areas. The intermediate segment does not change its shape notably and is thus already preorganized for cyclization (shown in grey). In contrast, the alignment of both terminal segments (shown in black) changes as the backbone kinks both Gly and Ser to sterically allow heptapeptide macrocycle formation. Both ends are predisposed to cyclization because of the hook-like conformation of the peptide.

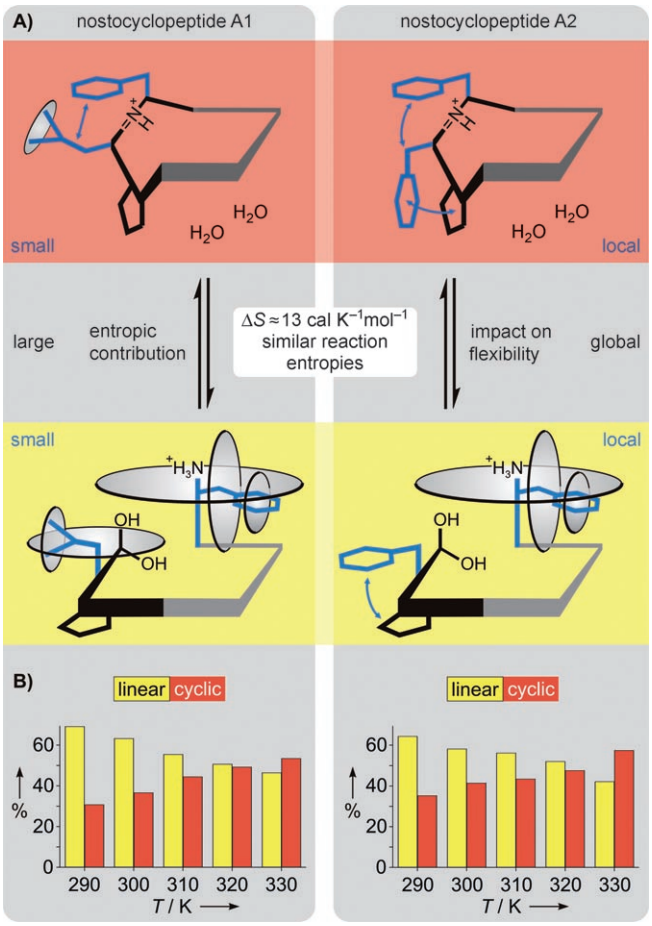

Figure 4. A) Depiction of cooperative stabilization in the peptide termini that result from hydrophobic side chain interactions in ncpA1 (left) and ncpA2 (right). Peptide segments and side chains that appear differently in the linear and cyclic peptides and in ncpA1 and ncpA2 are marked blue. The presence of NOE contacts is indicated by blue arrows, while the presence of rotational mobility is indicated by grey disks. To obtain quantitative pictures of the side chain mobilities, scalar coupling constants were used to calculate the respective rotamer populations. For example, in the case of rotation of the Tyr side chain around the C_α–C_β bond in ncpA1, equal populations of all three rotamers (*gauche–trans*, *trans–gauche*, *gauche–gauche*) in the linear peptide indicate unhindered side chain mobility; this is in contrast to the 0:78:22 population obtained for the cyclic species. The Leu side chain exhibits complicated spin systems, which prevent the calculation of rotamer populations. However, the complete absence of sequential NOE contacts (linear peptide) in contrast to the large number of dipolar couplings towards the Tyr side chain (cyclic peptide) give a qualitative picture of the interactions and mobilities displayed. B) Diagrams showing the temperature dependencies of the ncpA1 (left) and ncpA2 (right) macrocyclization equilibria. In spite of the differences in preorientation, the macrocyclization in both cases involves a similar entropy gain of ca. 13 cal K^{-1} mol^{-1}, which reveals that the entropic effects of local nonpolar interactions are small as compared to the cyclization entropy change.

temperature dependence of the ring-chain isomerism (Figure 4B) that is the key for obtaining a quantitative insight into the cyclization process. At pH 5.2, each peptide was allowed to equilibrate at the respective temperature (10 K steps between 290 and 330 K), and ^{1}H NMR spectra were subsequently recorded. The percentages of linear and cyclic peptide could be read out from these spectra by comparing corresponding ^{1}H NMR signal integrals. In spite of different local dynamics, both peptides exhibit almost identical cyclization behaviour (Figure 4B and the Supporting Information). With increasing temperature, the amount of cyclopeptides present for both grew almost at the same rate and in a near linear manner, yielding approximately 6% more cyclopeptide every 10 K. In both cases, a 1:1 mixture of linear and cyclic species was obtained at about 320 K.

The data displayed in Figure 4B allow for quantification of the change in dynamics as correlated with the thermodynamic parameter of entropy. Van't Hoff plots of the identical gradients yield cyclization enthalpies of approximately 4 kcal mol^{-1}, and the Gibbs–Helmholtz equation affords the entropy changes; for both peptides the macrocyclization is accompanied by an entropy gain of approximately 13 cal K^{-1} mol^{-1} (Supporting Information). A question of interest concerns the net entropy balance of the peptides upon macrocyclization, which in principle can also be assumed to be positive. In the case of the nostocyclopeptides, however, NMR coupling constants suggest that the cyclizing peptide chains loose entropy as a greater number of time-averaged coupling constants are observed after cyclization; this gives sound evidence for less conformational freedom after cyclization. Consequently, the positive entropy balance results from the release of water during the cyclization, which compensates for the entropically unfavourable macrocyclization step. The overall entropy is comprised of various contributions.[22] The hydration state is a complex factor which can significantly affect the total entropy balance.[23,24] In the case of the nostocyclopeptides, however, the observance of only minor chemical shift changes and structural calculations give evidence of negligible changes in hydration

upon macrocyclization, thus allowing the approximation of neglecting entropic and enthalpic contributions (Supporting Information). Since the amount and the entropy of water is the same for the cyclization process of each peptide, the loss in chain entropy must also be the same for both molecules, and must sum up to the experimental values of approximately 13 cal K^{-1} mol^{-1}. In other words, both chain-ring pairs exhibit roughly the same entropy difference in spite of the different conformational averaging of the linear peptides. The nearly identical cyclization behaviour of both peptides contrasts with the unequal preorientation of the linear cyclization precursors; this raises the question as to what extent a proper preorientation, which is well-known to promote macrocycle formation,[17,25] affects this process.

Our results indicate that side chain interactions operate on local segments yet do not measurably influence the overall backbone mobility. The covalent tethering of both termini is essential to constrict the global conformational space. The entropy of macrocyclization mainly depends on chain length and is largely independent of peptide conformation. This means that the mutually contradictory models of Flory's isolated-pair hypothesis[26] and the worm-like behaviour of polymer chains,[27] are both right when it comes to macrocyclization. The exceptionally selective ring closure of the nostocyclopeptides illustrates the importance of preorientation for preventing polymerization or other side reactions, while it only plays a minor role for the entropy balance. Expressed in a more quantitative way: the macroscopic property of rigidity can be transferred to the molecular level as the loss of rotational degrees of freedom of the peptide backbone—as linear conformations—are excluded. Linear conformations are excluded as rotational degrees of freedom are lost from the peptide backbone. The twelve independent ϕ and ψ angles of the linear heptapeptides (Pro and amine not included) become dependent within the cyclic heptapeptides and approximately six rotational degrees of freedom are lost. As a rule of thumb we can assume that cyclization divides the number of rotational degrees of freedom by a factor of two because no torsion angle can be rotated without the counterrotation of a second torsional angle of the peptide backbone. The entropy change accompanying the restriction of one single rotational degree of freedom is approximately -3 cal K^{-1} mol^{-1}.[28] Multiplied by 6, we expect a negative cyclization entropy of -18 cal K^{-1} mol^{-1} for the ring closures of ncpA1 and ncpA2. The positive experimental value of $+13$ cal K^{-1} mol^{-1} differs by 31 cal K^{-1} mol^{-1}; this fits amazingly well with the positive entropy of the two released water molecules[29,30] observed during imine formation. However, investigation of analogous macrocyclization systems is required to confirm the further validity of this correlation.

The reversible macrocyclic imine closure of the nostocyclopeptides has allowed for the quantification of the entropy balance of a biomolecule's macrocyclization, and has provided insight into the involved conformational restriction. The nearly identical cyclization behaviour of both peptides suggests that preorientation, while a prerequisite for efficient macrocycle formation, does not promote the cyclization process by lowering the loss of dynamics. Consequently, the preorientation of a peptide chain and its subsequent cyclization determine mobility and therewith entropy on different scales, and a well-defined solution conformation as it is observed by spectroscopic methods should not be mistaken for rigidity. Although macroscopic terms like "mobility" or "rigidity" have only limited applicability on the molecular level, they are descriptive for the complex concept of "conformational space". If a conclusion about proteins is to be drawn from the information obtained about the nostocyclopeptides, then it is that the linear peptides are reminiscent of the molten globule state of proteins, which is highly dynamic in spite of the presence of numerous secondary structural elements.[31] In contrast, protein surface loops may be unstructured, but as they are fixed at the positions emerging from the stiffened membrane-bound protein segment they can perform important tasks. This recalls an analogy to the cyclopeptides which, even if they exhibit conformational averaging, experience a considerable restriction of mobility due to the covalent tethering of both termini. The results obtained by this work demonstrate the importance and ubiquity of biological macrocycles, supporting the concept that linear substrates can populate preferred conformations, but that only cyclization generates selective ligands for the interaction with the dedicated molecular receptors.

Acknowledgements

We acknowledge financial support by the Deutsche Forschungsgemeinschaft, the Fonds der Chemischen Industrie, and the Studienstiftung des Deutschen Volkes (S.E.).

Keywords: imido compounds · macrocycles · NMR spectroscopy · peptides · structure elucidation

[1] F. Kopp, M. A. Marahiel, *Nat. Prod. Rep.* **2007**, *24*, 735–749.
[2] J. W. Trauger, R. M. Kohli, H. D. Mootz, M. A. Marahiel, C. T. Walsh, *Nature* **2000**, *407*, 215–218.
[3] C. T. Walsh, *Science* **2004**, *303*, 1805–1810.
[4] H. Kessler, *Angew. Chem.* **1982**, *94*, 509–520; *Angew. Chem., Int. Ed. Engl.* **1982**, *21*, 512–523.
[5] J. Rizo, L. M. Gierasch, *Annu. Rev. Biochem.* **1992**, *61*, 387–418.
[6] D. F. Veber, F. W. Holly, W. J. Paleveda, R. F. Nutt, S. J. Bergstrand, M. Torchiana, M. S. Glitzer, R. Saperstein, R. Hirschmann, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 2636–2640.
[7] J. Clardy, C. T. Walsh, *Nature* **2004**, *432*, 829–837.
[8] G. Weckbecker, I. Lewis, R. Albert, H. A. Schmid, D. Hoyer, C. Bruns, *Nat. Rev. Drug Discovery* **2003**, *2*, 999–1017.
[9] J. W. Smith, *Curr. Opin. Invest. Drugs (Thomson Curr. Drugs)* **2003**, *4*, 741–745.
[10] M. A. Dechantsreiter, E. Planker, B. Mathä, E. Lohof, G. Hölzemann, A. Jonczyk, S. L. Goodman, H. Kessler, *J. Med. Chem.* **1999**, *42*, 3033–3040.
[11] R. B. Silverman, *The Organic Chemistry of Drug Design and Drug Action*, Elsevier Academic Press, Oxford, **2004**.
[12] T. Golakoti, W. Y. Yoshida, S. Chaganty, R. E. Moore, *J. Nat. Prod.* **2001**, *64*, 54–59.
[13] J. E. Becker, R. E. Moore, B. S. Moore, *Gene* **2004**, *325*, 35–42.
[14] F. Kopp, C. Mahlert, J. Grünewald, M. A. Marahiel, *J. Am. Chem. Soc.* **2006**, *128*, 16478–16479.
[15] R. W. Layer, *Chem. Rev.* **1963**, *63*, 489–510.
[16] E. H. Cordes, W. P. Jencks, *J. Am. Chem. Soc.* **1962**, *84*, 832–837.

[17] S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders, J. F. Stoddart, *Angew. Chem.* **2002**, *114*, 938–993; *Angew. Chem. Int. Ed.* **2002**, *41*, 898–952.
[18] J.-M. Lehn, *Chem. Eur. J.* **1999**, *5*, 2455–2463.
[19] N. E. Borisova, M. D. Reshetova, Y. A. Ustynyuk, *Chem. Rev.* **2007**, *107*, 46–79.
[20] R. Cacciapaglia, S. Di Stefano, L. Mandolini, *J. Am. Chem. Soc.* **2005**, *127*, 13666–13671.
[21] G. Melacini, Q. Zhu, M. Goodman, *Biochemistry* **1997**, *36*, 1233–1241.
[22] K. K. Frederick, M. S. Marlow, K. G. Valentine, A. J. Wand, *Nature* **2007**, *448*, 325–329.
[23] H. Gohlke, G. Klebe, *Angew. Chem.* **2002**, *114*, 2764–2798; *Angew. Chem. Int. Ed.* **2002**, *41*, 2644–2676.
[24] J. Dolenc, R. Baron, J. H. Missimer, M. O. Steinmetz, W. F. van Gunsteren, *ChemBioChem* **2008**, *9*, 1749–1756.
[25] J. Blankenstein, J. Zhu, *Eur. J. Org. Chem.* **2005**, 1949–1964.
[26] P. J. Flory, *Statistical Mechanics of Chain Molecules*, Wiley-Interscience, New York, **1969**.
[27] Y. Z. Ohkubo, C. L. Brooks III, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 13916–13921.
[28] M. Mammen, E. I. Shakhnovich, G. M. Whitesides, *J. Org. Chem.* **1998**, *63*, 3168–3175.
[29] L. M. Amzel, *Proteins: Struct., Funct., Genet.* **1997**, *28*, 144–149.
[30] B. Lu, C. F. Wong, *Biopolymers* **2005**, *79*, 277–285.
[31] S. Hiller, G. Wider, L. L. Imbach, K. Wüthrich, *Angew. Chem.* **2008**, *120*, 992–996; *Angew. Chem. Int. Ed.* **2008**, *47*, 977–981.

Received: May 7, 2008
Published online on September 26, 2008

